



Computing Summary

F. Bianchi
Torino

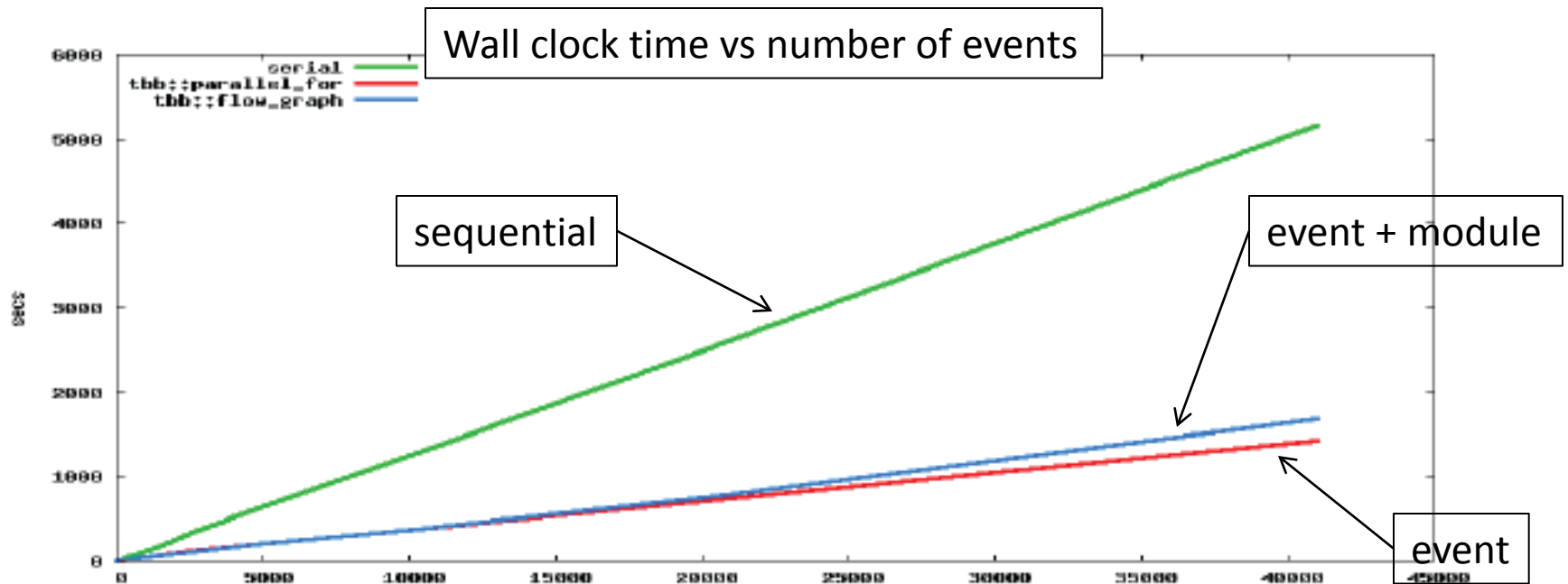
V SuperB Collaboration Meeting
Pisa, September 21st, 2012

Outline

- Lively parallel sessions!
 - R&D
 - Distributed Computing Tools
 - FullSim & Background
 - Computing section of Detector TDR
- My selection of presented material plus some comments
 - Look at the presentations in Indico for details

R&D: Parallel Framework (1)

- M. Corvo has presented nice progresses.
- Test were made using the current FastSim Framework that is based on the BaBar code.
- Test both of event level and event plus module parallelism.



R&D: Parallel Framework (2)

- We need to come to a full understanding of all dependencies among modules
- This is a good starting point, but there are a lot of analysis patterns to consider, not to mention a full reconstruction framework
- This exercise was helpful to understand also what modules can do and what modules must NOT do to improve optimization
- There's room for improvements also in the generation part, trying to limit (or at least factorize) the use of Common Blocks in Fortran code
- In the long term we will abandon the current framework for a new one which is natively parallel and whose architecture will be designed according to our experiences

R&D: Distributed Storage

- Storage system evaluation
 - HadoopFS on WAN: testbed on Bari and Napoli
- Data access framework library development
 - Data access optimization on local and WAN scenario
 - Mask the low level data access layer at the sites
 - Useful support from ROOT development team
- Mass data transfer system
 - FTS3 evaluation
 - PhEDEx evaluation process (standby)
- File cataloging (dynamic LFCing by EMI R&D, standby)

DIRAC (1)

- DIRAC (*Distributed Infrastructure with Remote Agent Control*) is a Grid resource management tool.
 - Goal: setup and configure a Dirac system to fulfil the SuperB requirements
 - General work plan:
 - Simulation production use case (in progress)
 - Workload Monitor system
 - Analysis use case integrated with Ganga system
 - Mass data transfer system

DIRAC (2)

- Present goal: manage all aspects related to Simulation Production (FastSim and FullSim) via DIRAC
 - User management (role and permissions)
 - Site Management per Session
 - Production creation and monitoring
 - Requests creation and monitoring
 - Bunch jobs submission and monitoring
 - DIRAC need to interact with Bookkeeping Database (SBK5)
- Test to launch Severus (python wrapper for SuperB simulation production) via DIRAC have been successfully performed
- Very productive interactions with the DIRAC development team

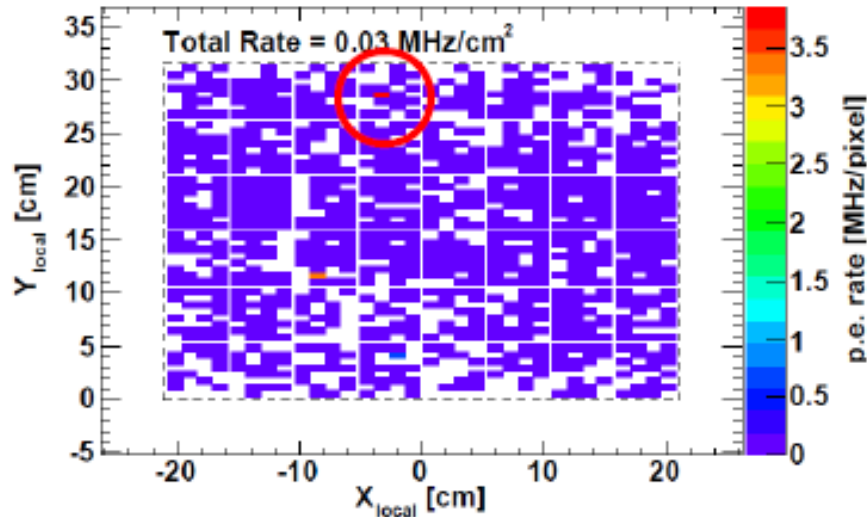
FullSim & Background

- Summer production was a success
 - A very complete set of background samples have been produced
 - Rad-bhabha (low and high k)
 - Pairs
 - Touschek and BeamGas (HER/LER)
 - Synchrotron Radiation (HER/LER)
 - Several improvement in detector and B field model
- Production was affected by a bug on optical photons in FDIRC
 - Luckily was recoverable offline
 - Need to improve our QA protocol before starting production
 - More advance planning of production is needed
- Good stress test of new distributed production system
 - Transfer of large (>3 GeV) files over the Grid failed.

The patch in action

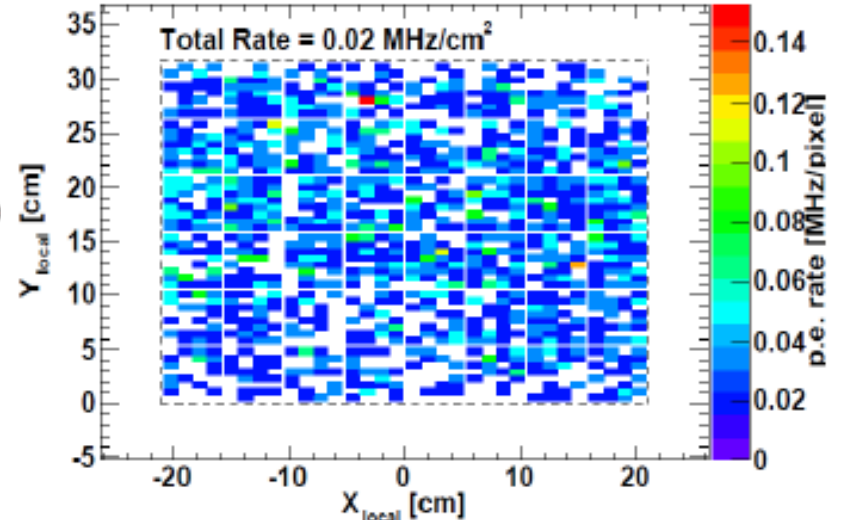
Before the patch

Total Photo-electron rates on FDIRC sector 0

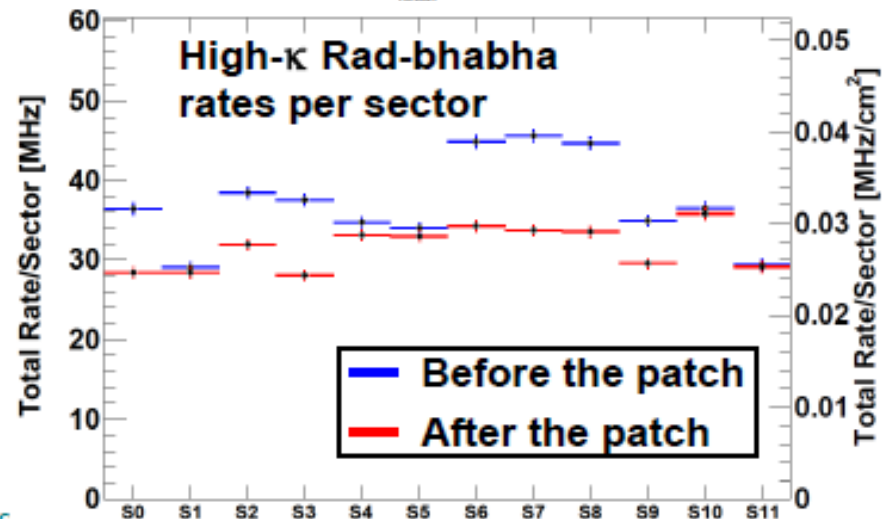


After the patch

Total Photo-electron rates on FDIRC sector 0 (corrected)



- Algorithm correctly identify “hot pixels” and filter them out
- The total rates in some sectors get significantly reduced by up to 25-30%



Next productions

- Currently we are exploiting resources on the Grid for central production(s).
 - But output is sent back to CNAF and analyzed there.
- We should start to use resources on the Grid also for analysis.
- We have a Ganga based tool to submit analysis job on the Grid.
 - It has been tested by few volunteers
 - Should become of general use sometime early next year.
 - Our plan is to offer a tutorial at the first 2013 meeting and ask people to use it for their analysis soon after.

TDR Writing Recommendation

- Please remember that the repository is case sensitive.
 - The files "myfile.pdf" and "myfile.PDF" are different files.
- Please avoid committing to the repository a "dtdr.tex" file with all subsystems (with the exclusion of yours) commented out.
- Please resolve all conflicts before committing the text to the repository
 - Do a "svn update"
 - Resolve all conflicts by editing the affected files
 - Finally do the "svn commit"
- Please avoid lines with >1024 characters
- Check with the "svn stat" command
 - Files marked with a "c" have conflicts
 - Files marked with a "?" are missing

Computing Section of DTDR

Status

- All material is there
 - Only final summary is missing
- More work by the computing group to improve wording is needed.
 - Plan is to complete this step in the next couple of weeks

14 Software and Computing		339
14.1 Computing Overview	F.Bianchi 2 pages	339
14.2 Tools to support detector studies	F.Bianchi 1 pages	339
14.2.1 Full Simulation	A. Di Simone - E. Paoloni - A. Perez 4 pages	339
14.2.1.1 Bruno: the SuperB full simulation software		340
14.2.1.2 Geometry description		340
14.2.1.3 Simulation input: Event generators		340
14.2.1.4 Simulation output: Hits and MonteCarlo Truth		341
14.2.1.5 Simulation optimization		341
14.2.1.6 Staged simulation		341
14.2.1.7 Interplay with fast simulation		342
14.2.1.8 Long term evolution of the full simulation software		342
14.2.2 Fast Simulation	M. Rama 4 pages	343
14.2.2.1 Event generation		343
14.2.2.2 Detector description		344
14.2.2.3 Interaction of particles with matter		344
14.2.2.4 Detector response		344
14.2.2.5 Reconstruction		345
14.2.2.6 Machine backgrounds		345
14.2.2.7 Analysis tools		346
14.2.2.8 Simulation validation and detector studies		347

	14.2.2.5	EMBEDDED VIRTUALIZATION AND DISTANCE SERVICES	344
14.2.3		Distributed computing tools G. Donvito - A. Fella - E. Luppi - S. Pardi L. Tomassetti 10 pages	347
	14.2.3.1	Distributed resources	348
	14.2.3.2	Distributed systems design: a bird's-eye view	349
	14.2.3.3	The production system	349
	14.2.3.4	The data analysis system prototype	351
	14.2.3.5	The bookkeeping and data placement database	352
14.2.4		Collaborative tools M. Corvo - A. Gianoli - S. Longo - R. Stroili 2 pages	353
	14.2.4.1	Overview	353
	14.2.4.2	Authorization	353
	14.2.4.3	Portal System	353
	14.2.4.4	Document repository	353
	14.2.4.5	Documentation	354
	14.2.4.6	Code repository	354
	14.2.4.7	Code packaging and distribution	355

14.3	Computing model outline	F. Bianchi - A. Fella - C. Grandi - S. Luitz - E. Luppi - S. Pardi - L. Tomassetti	6 pages	356
14.3.1	Data processing			356
14.3.2	Resource estimate	F. Bianchi - S. Luitz	4 pages	357
14.3.3	Computing Infrastructure	F. Bianchi - S. Luitz - S. Pardi	4 pages	357
14.4	R & D program	M. Corvo - G. Donvito - A. Fella - F. Giacomini - S. Longo - S. Pardi	8 pages	358
14.4.1	R& D on parallelization			358
14.4.2	GPU R& D			359
14.4.3	Framework R & D			360
14.4.4	DIRAC framework evaluation			362
14.4.4.1	Pilot jobs model			363
14.4.4.2	Dirac data management			363
14.4.4.3	DIRAC API			364
14.4.4.4	User Management			364
14.4.4.5	Tested Use Cases			364
14.4.4.6	SuperB DIRAC module			364
14.4.4.7	Building up a DIRAC Infrastructure for SuperB			364
14.4.4.8	Future Works			365
14.4.5	Data management and distributed storage R&D			365
14.4.5.1	WAN data access			365
14.4.5.2	Data access library			366
14.4.5.3	File Transfer Service evolution			366
14.4.5.4	Dynamic file catalogue technology			366
14.4.5.5	Storage system evaluation			367
14.5	Summary	F. Bianchi	1 pages	367

Conclusions

- The computing group is supporting the collaboration by providing:
 - Collaborative Tools
 - Physics Tools: FastSim, etc.
 - FullSim
 - Production Tools
 - Bookkeeping Tools
- There is an active R&D program aimed at the design of the computing model.
- The activities funded under the Pon ReCaS are an important step forward into building the computing infrastructure.
- A severe lack of manpower is affecting us.
- Come and join the fun !